

31759-190545
Akihiro OKUMURA et al.

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2002年 6月27日

出願番号

Application Number:

特願2002-187625

[ST.10/C]:

[J P 2002-187625]

出願人

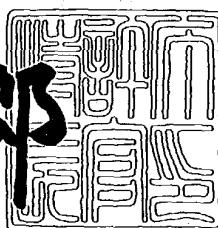
Applicant(s):

沖電気工業株式会社

2003年 4月 1日

特許庁長官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3022441

【書類名】 特許願

【整理番号】 KN002521

【提出日】 平成14年 6月27日

【あて先】 特許庁長官 及川 耕造 殿

【国際特許分類】 G06F 15/00

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

【氏名】 松村 靖子

【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

【氏名】 池野 篤司

【特許出願人】

【識別番号】 000000295

【氏名又は名称】 沖電気工業株式会社

【代表者】 篠塚 勝正

【代理人】

【識別番号】 100090620

【弁理士】

【氏名又は名称】 工藤 宣幸

【手数料の表示】

【予納台帳番号】 013664

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9006358

特2002-187625

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 主要箇所判定装置

【特許請求の範囲】

【請求項1】 領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する装置において、

制御文字を解析して表示領域を抽出する領域抽出部と、

抽出された表示領域の大きさを比較し、比較した結果、最も大きい領域が主要箇所であると判定する判定部と

を備えることを特徴とする主要箇所判定装置。

【請求項2】 領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する装置において、

制御文字を解析して表示領域を抽出する領域抽出部と、

抽出された表示領域の位置を比較し、比較した結果、最も中心に近い領域が主要箇所であると判定する判定部と

を備えることを特徴とする主要箇所判定装置。

【請求項3】 領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する装置において、

制御文字を解析して表示領域を抽出する領域抽出部と、

抽出された表示領域の大きさを比較し、比較した結果、最も大きい領域を選択する最大領域判定部と、

抽出された表示領域の位置を比較し、比較した結果、最も中心に近い領域を選択する中心領域判定部と、

最も大きい領域、または、最も中心に近い領域のどちらかを主要箇所であると判定する判定部と

を備えることを特徴とする主要箇所判定装置。

【請求項4】 上記判定部は、(1)式を用いて計算した値を比較することによって、最も大きい領域、または、最も中心に近い領域のどちらかを主要箇所であると判定することを特徴とする請求項3に記載の主要箇所判定装置。

$$A = X \alpha + Y \beta \quad (1)$$

ここで、Xは抽出された領域の面積、Yは表示される文書全体の中心位置と抽出された領域の中心位置との距離の逆数を表し、 α 、 β は、重み係数である。

【請求項5】 上記判定部は、(2)式を用いて計算した値を比較することによって、最も大きい領域、または、最も中心に近い領域のどちらかを主要箇所であると判定することを特徴とする請求項3に記載の主要箇所判定装置。

$$A = X Y \quad (2)$$

ここで、Xは抽出された領域の面積、Yは表示される文書全体の中心位置と抽出された領域の中心位置との距離の逆数を表す。

【請求項6】 上記判定部は、最も大きい領域、または、最も中心に近い領域のどちらかを主要箇所であると判定するにつき、各領域のそれぞれで、同種同サイズの文字の個数を計数し、その個数が多いほうを主要箇所であると判定することを特徴とする請求項3に記載の主要箇所判定装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する主要箇所判定装置に関するものである。

【0002】

【従来の技術】

WWW (W o r l d W i d e W e b) サイトに存在するWWW文書を取得し閲覧するためのツールとして、WWWブラウザがある。一般的に、WWW文書は、HTMLのような言語を用いて書かれており、その文書のページのレイアウト、文字の大きさなどを柔軟に指定することができるようになっている。特に、図1のように、タイトル(A)、他のリンク先(B)、本文(C)、その他(D)、(E)など、ページがいくつかの領域に分割されて、WWWブラウザに表示されるようなWWW文書が多く見られる。WWWブラウザを用いて、このようなWWW文書から必要な情報を得るために、ユーザは、目的のWWW文書のURLを指定し、その文書がWWWブラウザ上に表示された後に、文書をスクロール

しながら目視により検索したり（人手による検索）、あるいは文字列検索機能を利用するといった作業を行なう必要がある。

【0003】

例えば、図1の（C）の部分が、ユーザの必要とする文書であったとし、こういったWWWページが多数ある場合には、そのユーザが必要とする情報のみを複数のWWWページから自動的にスクラップし、1つの文書にまとめてユーザに提示されることが、人手による作業を簡略化する上で望ましくなる。このようなWWW情報抽出システムが、特開平10-187753号公報に示されている。

【0004】

【発明が解決しようとする課題】

しかしながら、上記におけるWWW情報抽出システムでは、ユーザがWWW文書中で自分が必要とするデータの開始箇所と終了箇所をあらかじめ手入力により指定することが必要であり、領域分割の形が複数種類あった場合には、人手による作業量が増大してしまう。また、新しい領域分割の形のページを対象にしようとすると、さらに入手の作業が必要となる。

【0005】

この必要とするデータの箇所を自動的に判定できることが望ましい。また、自動的に判定できることにより、（イ）指定ウェブページが更新されたときにユーザに通知する、（ロ）検索（主要箇所以外は検索対象としない、など）、（ハ）要約（主要箇所のみを要約対象とする、など）などのサービスやシステムを容易に構築することが可能となる。

【0006】

そのため、WWW文書等の文書から、ユーザが必要とする箇所を自動的に判定する主要箇所判定装置が望まれている。

【0007】

【課題を解決するための手段】

第1の本発明は、領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する主要箇所判定装置において、制御文字を解析して表示領域を抽出する領域抽出部と、抽出された表示領域の大きさを比較し、比

較した結果、最も大きい領域が主要箇所であると判定する判定部とを備えることを特徴とする。

【0008】

第1の本発明は、領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する主要箇所判定装置において、制御文字を解析して表示領域を抽出する領域抽出部と、抽出された表示領域の位置を比較し、比較した結果、最も中心に近い領域が主要箇所であると判定する判定部とを備えることを特徴とする。

【0009】

第3の本発明は、領域が分割されて表示されるドキュメントに対して、ユーザが必要とする主要箇所を判定する主要箇所判定装置において、制御文字を解析して表示領域を抽出する領域抽出部と、抽出された表示領域の大きさを比較し、比較した結果、最も大きい領域を選択する最大領域判定部と、抽出された表示領域の位置を比較し、比較した結果、最も中心に近い領域を選択する中心領域判定部と、最も大きい領域、または、最も中心に近い領域のどちらかを主要箇所であると判定する判定部とを備えることを特徴とする。

【0010】

【発明の実施の形態】

(A) 第1の実施形態

(A-1) 第1の実施形態の構成

図2は、本発明による主要箇所判定装置の第1の実施形態の機能的構成を示すブロック図である。なお、第1の実施形態の主要箇所判定装置は、パソコン等の情報処理装置で実現されるが、機能的には、図2で表すことができる。

【0011】

端子11は入力端子である。入力は、図1のようにWWWブラウザに表示されるWWW文書のソースであり、実際には、図3にあるように、HTMLのような言語で記述されているようなデータである。

【0012】

領域抽出部101は、入力されたWWW文書のソースの中にある、領域分割に

関する制御文字を解析することにより、分割された領域を抽出し、それぞれの分割された領域の表示される面積を計算する。領域抽出部101は、それぞれの分割された領域の表示面積と、その領域に表示されるソース部分を分割領域個数分だけ出力する。出力する順番は、例えば、右の分割領域から左へ、上の分割領域から下へのラスタースキャンの順序であるとする。

【0013】

最大領域判定部102は、1つまたは複数の面積とWWW文書のソースの組を受け取り、それらのデータの組のうち、面積が最大値をとるデータの組に対応するWWW文書のソースを出力する。

【0014】

端子12は出力端子である。

【0015】

(A-2) 第1の実施形態の動作

動作の概要としては、入力端子11より、WWW文書が入力され、その入力WWW文書の内、一番大きくWWWブラウザに表示される領域のWWW文書（のソース）を出力端子12より出力する。

【0016】

例えば、図1で表示されるようなWWW文書のソース、すなわち図3のようなデータが入力端子11より入力された場合、領域抽出部101では、例えば、図4のような出力結果となる。出力結果の1レコードは、（分割領域の面積、分割領域のソース部分）を表す。

【0017】

図3のようなHTMLで記述された場合、<FRAMESET ROWS="20%, 80%">と記述された部分が、領域分割を示す制御文字の一つである。上記の例の場合、表示部分を行方向に（すなわち上下に）20%と80%の割合で領域分割するという意味であり、このように記述されている分割比率などを基に面積を計算する。面積は、その領域が占める画素数で表してもよいし、図4での例のように、全体を100とした面積の割合でもよい。また、通常の面積の単位（平方ミリメートルや平方センチメートルなど）で記述してもよい。

【0018】

最大領域判定部102は、領域抽出部101より入力された分割領域の面積を参照し、面積が最大の分割領域のソース部分を出力する。

【0019】

最終的に、出力端子12より、入力された文書のソースから、面積が最大の分割領域のソース部分を出力する。

【0020】

例えば、図3のようなデータが入力端子11より入力される場合、最終的に、図5のようなデータが出力される。

【0021】

(A-3) 第1の実施形態の効果

以上のように、第1の実施形態によれば、領域が分割されて表示されるドキュメントに対して、それらドキュメントの主要部分は、領域の面積が最大となることが多いという性質を利用して、ユーザが必要とする箇所を、その面積が最大となる領域と判定することにより、自動化が可能で、入手による作業を省略することができる。また、単なる面積の値の比較だけなので、計算が非常に簡単であり、高速に結果を得ることができる。

【0022】

また、この装置は、自動化し、簡単かつ高速にユーザが必要とする箇所を特定できるため、(イ) 指定ウェブページのうち必要箇所が更新されたときにユーザに通知する、(ロ) 検索(必要箇所以外は検索対象としない、など)、(ハ) 要約(必要箇所のみを要約対象とする、など)などの部分的手段または部品として利用することが可能となる。

【0023】

(B) 第2の実施形態

(B-1) 第2の実施形態の構成

図6は、本発明による主要箇所判定装置の第1の実施形態の機能的構成を示すブロック図である。

【0024】

端子21は入力端子である。入力は、図1のようにWWWブラウザに表示されるWWW文書のソースであり、実際には、図3にあるように、HTMLのような言語で記述されているようなデータである。

【0025】

領域抽出部201は、第1の実施形態における領域抽出部101とは少し異なり、入力されたWWW文書のソースの中にある、領域分割に関する制御文字を解析することにより、分割された領域を抽出し、それぞれの分割された領域の表示される位置を計算する。領域抽出部201は、それぞれの分割された領域の位置情報と、その領域に表示されるソース部分を分割領域個数分だけ出力する。出力する順番は、例えば、右の分割領域から左へ、上の分割領域から下へのラスタースキャンの順序であるとする。

【0026】

中心領域判定部202は、1つまたは複数の領域位置情報とWWW文書のソースの組を受け取り、それらのデータの組のうち、表示される文書の中心に最も近い位置情報をもつデータの組に対応するWWW文書のソースを出力する。

【0027】

端子22は出力端子である。

【0028】

(B-2) 第2の実施形態の動作

動作の概要としては、入力端子21より、WWW文書が入力され、その入力WWW文書の内、最も中心に近い位置に表示される領域のWWW文書（のソース）を出力端子22より出力する。

【0029】

例えば、図1で表示されるようなWWW文書のソース、すなわち図3のようなデータが入力端子21より入力された場合、領域抽出部201では、例えば、図7のような出力結果となる。出力結果の1レコードは、（分割領域の左上の位置座標、分割領域の右下の位置座標、分割領域のソース部分）を表す。位置座標は、最左上を（0, 0）とし、左方向もしくは下方向へ離れる位置を正值とする。また、位置座標は、（左方向へ離れる距離、下方向へ離れる距離）と表すことに

する。分割領域の左上の位置座標および分割領域の右下の位置座標が、この分割領域の位置情報となる。

【0030】

図3のようなHTMLで記述された場合、<FRAMESET ROWS="20%, 80%">と記述された部分が、領域分割を示す制御文字の一つである。上記の例の場合、表示部分を行方向に（すなわち上下に）20%と80%の割合で領域分割するという意味であり、このように記述されている分割比率などを基に位置情報を計算する。位置座標の単位（最左上からの距離の単位）は、その距離を画素数で表示してもよいし、図7での例のように全体の縦方向の長さおよび横方向の長さを100とした時の距離で表してもよい。また、通常の長さ（距離）の単位（mmやcmなど）で記述してもよい。

【0031】

中心領域判定部202は、領域抽出部201より入力された分割領域の位置情報を参照し、WWW文書の表示される中心点（図1のように表示される文書で、領域抽出部201からの出力が図7のような例の場合、中心点は（50、50））を含む分割領域のソース部分を出力する。

【0032】

最終的に、出力端子22より、入力された文書のソースから、最も中心に近い位置に表示される分割領域のソース部分を出力する。

【0033】

例えば、図3のようなデータが端子21より入力される場合、最終的に、図5のようなデータが出力される。

【0034】

(B-3) 第2の実施形態の効果

以上のように、第2の実施形態によれば、領域が分割されて表示されるドキュメントに対して、それらドキュメントの主要部分は、領域の中心部分に表示されることが多いという性質を利用して、ユーザが必要とする箇所を、最も中心に近い位置に表示される領域と判定することにより、自動化が可能で、人手による作業を省略することができる。また、単なる表示位置情報の比較だけなので、計算

が非常に簡単であり、高速に結果を得ることができる。

【0035】

また、この装置は、自動化し、簡単かつ高速にユーザが必要とする箇所を特定できるため、（イ）指定ウェブページのうち必要箇所が更新されたときにユーザに通知する、（ロ）検索（必要箇所以外は検索対象としない、など）、（ハ）要約（必要箇所のみを要約対象とする、など）などの部分的手段または部品として利用することが可能となる。

【0036】

(C) 第3の実施形態

(C-1) 第3の実施形態の構成

図8は、本発明による主要箇所判定装置の第3の実施形態の機能的構成を示すブロック図である。

【0037】

端子31は入力端子である。入力は、第1、2の実施形態と同様に、図1のようにWWWブラウザに表示されるWWW文書のソースであり、実際には、図3にあるように、HTMLのような言語で記述されているようなデータである。

【0038】

領域抽出部301は、第1の実施形態の領域抽出部101と第2の実施形態の領域抽出部201の2つの動作を同時に行うような部分であり、入力されたWWW文書のソースの中にある、領域分割に関する制御文字を解析することにより、分割された領域を抽出し、それぞれの分割された領域の表示される面積を計算すると同時に、それぞれの分割された領域の表示される位置を計算する。領域抽出部301は、それぞれの分割された領域の表示面積、および、それぞれの分割された領域の位置情報、その領域に表示されるソース部分を分割領域個数分だけ出力する。出力する順番は、例えば、右の分割領域から左へ、上の分割領域から下へのラスタースキャンの順序であるとする。

【0039】

最大領域判定部302は、1つまたは複数の面積、位置情報、およびWWW文書のソースの組を受け取り、それらのデータの組のうち、面積が最大値をとる

データの組、すなわち、最大面積値、およびそれに対応する位置情報とWWW文書のソースを出力する。

【0040】

また、中心領域判定部3022は、1つまたは複数の面積、位置情報、およびWWW文書のソースの組を受け取り、それらのデータの組のうち、表示される文書の中心に最も近い位置情報をもつデータの組、すなわち、中心に最も近く表示される領域の面積、位置情報およびその領域に対応するWWW文書のソースを出力する。

【0041】

選択部303は、2つの領域に対する面積、位置情報、およびWWW文書のソースの組を受け取り、入力のどちらかを選択して、その選択されたデータの組のWWW文書のソースを出力する。

【0042】

端子32は出力端子である。

【0043】

(C-2) 第3の実施形態の動作

動作の概要としては、入力端子31より、WWW文書が入力され、その入力WWW文書の内、一番大きくWWWブラウザに表示される領域、または、最も中心に近い位置に表示される領域、のWWW文書（のソース）を後述する判定基準で選択し、出力端子32より出力する。

【0044】

例えば、図1で表示されるようなWWW文書のソース、すなわち図3のようなデータが入力端子31より入力された場合、領域抽出部301では、例えば、図9のような出力結果となる。この出力結果の1レコードは、（分割領域の面積、分割領域の左上の位置座標、分割領域の右下の位置座標、分割領域のソース部分）を表す。

【0045】

図3のようなHTMLで記述された場合、<FRAMESET ROWS="20%，80%">と記述された部分が、領域分割を示す制御文字の一つである

。上記の例の場合、表示部分を行方向に（すなわち上下に）20%と80%の割合で領域分割するという意味であり、このように記述されている分割比率などを基に面積および位置座標を計算する。

【0046】

面積は、その領域が占める画素数で表示してもよいし、図9での例のように全体を100とした面積の割合でもよい。また、通常の面積の単位で記述してもよい。

【0047】

位置座標は、最左上を(0, 0)とし、左方向もしくは下方向へ離れる位置を正値とする。また位置座標は、(左方向へ離れる距離、下方向へ離れる距離)と表すこととする。また、位置座標の単位(最左上からの距離の単位)は、その距離を画素数で表示してもよいし、図9での例のように全体の縦方向の長さおよび横方向の長さを100とした時の距離で表してもよい。また、通常の長さ(距離)の単位(mmやcmなど)で記述してもよい。第2の実施形態と同様に、分割領域の左上の位置座標および分割領域の右下の位置座標が、この分割領域の位置情報となる。

【0048】

最大領域判定部3021は、領域抽出部301より入力された分割領域の面積を参照し、面積が最大の分割領域の面積、位置情報、およびソース部分を出力する。

【0049】

また、中心領域判定部3022は、領域抽出部301より入力された分割領域の位置情報を参照し、WWW文書の表示される中心点(図1のように表示される文書で、領域抽出部301からの出力が図9のような例の場合、中心点は(50, 50))を含む分割領域の、面積、位置情報、およびソース部分を出力する。

【0050】

選択部303は、最大領域判定部3021から出力されたデータ、および、中心領域判定部3022から出力されたデータを受け取り、それぞれの入力データのうち、面積、および位置情報を用いて計算することによって、どちらかのデータ

タを選択する。

【0051】

例えば、次のような値Aをそれぞれの入力データに対して計算する。

【0052】

$$A = X\alpha + Y\beta \quad (1)$$

ここで、Xは面積、Yは表示される全体の中心位置と各分割領域の中心位置との距離の逆数を表す。α、βは、重み係数である。

【0053】

最大領域判定部3021から入力されるデータから(1)式で計算される値(Aとする)と、中心領域判定部3022から入力されるデータから(1)式で計算される値(Bとする)を比較し、AとBのうち大きいほうのデータを選択し、選択されたデータに対するWWW文書のソース部分を選択部303の出力とする

【0054】

例に挙げた場合は、最大領域判定部3021より出力された分割領域と中心領域判定部3022より出力された分割領域が一致しているので、選択部303より出力される分割領域は、最大領域判定部3021および中心領域判定部3022より出力される分割領域と同じになる。

【0055】

最終的に、端子32より、入力された文書のソースから、面積が最大、もしくは、最も中心に近い分割領域のうち、上記判定基準((1)式)に基づいて選択された領域のソース部分を出力することになる。

【0056】

(C-3) 第3の実施形態の効果

以上のように、第3の実施形態によれば、領域が分割されて表示されるドキュメントに対して、それらドキュメントの主要部分は、領域の面積が最大となる、または、領域の中心部分に表示されることが多いという性質を利用して、ユーザが必要とする箇所を、その面積が最大、または、中心点が含まれる領域と判定することにより、自動化が可能で、人手による作業を省略することができる。また

、単なる面積の値の比較だけなので、計算が非常に簡単であり、高速に結果を得ることができる。

【0057】

また、第1の実施形態と第2の実施形態の効果の両方を兼ね備えていることから、より正確にユーザの必要とする箇所を特定することができる。

【0058】

また、第1、第2の実施形態と同様に、この装置は、自動化し、簡単かつ高速にユーザが必要とする箇所を特定できるため、（イ）指定ウェブページのうち必要箇所が更新されたときにユーザに通知する、（ロ）検索（必要箇所以外は検索対象としない、など）、（ハ）要約（必要箇所のみを要約対象とする、など）などの部分的手段または部品として利用することが可能となる。

【0059】

(D) 他の実施形態

上記各実施形態では、例としてHTMLで記述された文書を例に挙げているが、HTMLだけでなくSGMLやXMLなどの、ページレイアウトなどの制御文字が含まれるような記述言語で記述された文書にも適用できる。

【0060】

第3の実施形態では、選択部303の判定基準として、（1）式にあるような計算式を用いた値を使っているが、

$$A = X Y \quad (2)$$

など、別の計算式を用いることもできる。

【0061】

また、第3の実施形態では、選択部303の判定基準として、面積と位置情報から計算された値を用いているが、別の情報を判断基準とすることもできる。例えば、領域に表示される同じ文字種かつ同サイズの文字数が多いほうを選択するといった判断基準を用いることもできる。

【0062】

また、上記各実施形態において、最大領域判定部や中心領域判定部の出力する順番は、右の分割領域から左へ、上の分割領域から下へのラスタースキャンの順

序としているが、その他の順序（「右の分割領域から左へ、上の分割領域から下へのラスタースキャンの順序」の逆、など）で出力することも可能である。

【0063】

また、上記各実施形態において、領域抽出部の出力として、図4、図7、図9のように、別ファイルへのポインタ（リンク）だけの場合を例として挙げているが、当該ファイルを取得してその内容を文書のソース部分として出力するものであってもよい。

【0064】

【発明の効果】

本発明によれば、WWW文書等の文書から、ユーザが必要とする箇所を自動的に判定する主要箇所判定装置を提供することができる。

【図面の簡単な説明】

【図1】

WWW文書の構成例の説明図である。

【図2】

第1の実施形態の主要箇所判定装置の機能的構成を示すブロック図である。

【図3】

第1の実施形態の主要箇所判定装置への入力データ例を示す説明図である。

【図4】

第1の実施形態の領域抽出部による抽出結果を示す説明図である。

【図5】

第1の実施形態の主要箇所判定装置の出力結果例を示す説明図である。

【図6】

第2の実施形態の主要箇所判定装置の機能的構成を示すブロック図である。

【図7】

第2の実施形態の領域抽出部による抽出結果を示す説明図である。

【図8】

第3の実施形態の主要箇所判定装置の機能的構成を示すブロック図である。

【図9】

第3の実施形態の領域抽出部による抽出結果を示す説明図である。

【符号の説明】

101、201、301…領域抽出部、

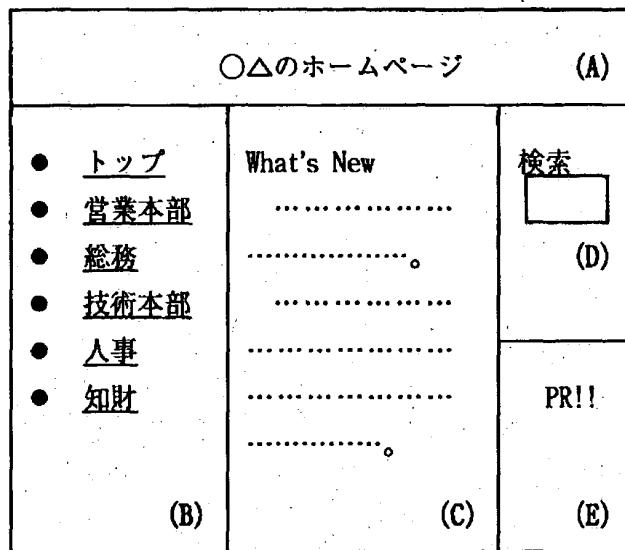
102、3021…最大領域判定部、

202、3022…中心領域判定部、

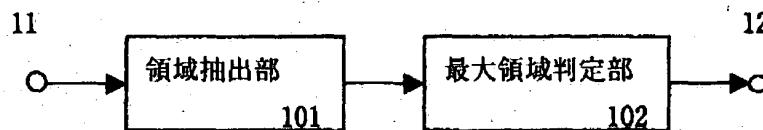
303…選択部。

【書類名】 図面

【図1】



【図2】



【図3】

```

<HTML>

<FRAMESET ROWS="20%,80%">
<FRAME SRC="title.html"> ←(A)の部分のソース
<FRAMESET COLS="30%,50%,20%">
<FRAME SRC="link.html"> ←(B)の部分のソース
<FRAME SRC="honbun.html"> ←(C)の部分のソース
<FRAMESET ROWS="50%,50%">
<FRAME SRC="sonota1.html"> ←(D)の部分のソース
<FRAME SRC="sonota2.html"> ←(E)の部分のソース
</FRAMESET>
</FRAMESET>
</FRAMESET>

</HTML>

```

【図4】

```

(20, <FRAME SRC="title.html">) ←(A)の部分
(24, <FRAME SRC="link.html">) ←(B)の部分
(40, <FRAME SRC="honbun.html">) ←(C)の部分
(8, <FRAME SRC="sonota1.html">) ←(D)の部分
(8, <FRAME SRC="sonota2.html">) ←(E)の部分

```

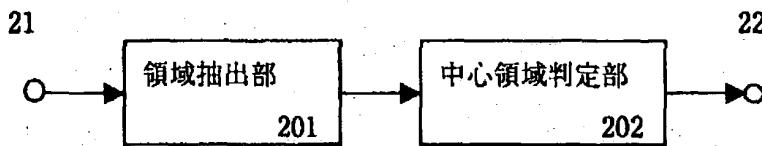
【図5】

```

<FRAME SRC="honbun.html">

```

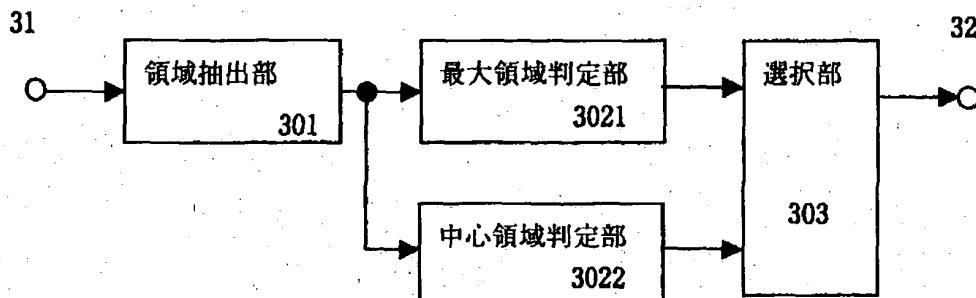
【図6】



【図7】

((0,0), (100,20), <FRAME SRC="title.html">)	←(A)の部分
((0,20), (30,100), <FRAME SRC="link.html">)	←(B)の部分
((30,20), (80,100), <FRAME SRC="honbun.html">)	←(C)の部分
((80,20), (100,60), <FRAME SRC="sonota1.html">)	←(D)の部分
((80,60), (100,100), <FRAME SRC="sonota2.html">)	←(E)の部分

【図8】



【図9】

(20, (0,0), (100,20), <FRAME SRC="title.html">)	←(A)の部分
(24, (0,20), (30,100), <FRAME SRC="link.html">)	←(B)の部分
(40, (30,20), (80,100), <FRAME SRC="honbun.html">)	←(C)の部分
(8, (80,20), (100,60), <FRAME SRC="sonota1.html">)	←(D)の部分
(8, (80,60), (100,100), <FRAME SRC="sonota2.html">)	←(E)の部分

【書類名】 要約書

【要約】

【課題】 WWW文書等の文書から、ユーザが必要とする箇所を自動的に判定する主要箇所判定装置を実現する。

【解決手段】 本発明の主要箇所判定装置は、文書中の制御文字を解析して表示領域を抽出する領域抽出部と、抽出された表示領域の大きさを比較し、比較した結果、最も大きい領域が主要箇所であると判定する判定部とを備える。判定部は、抽出された表示領域の位置を比較し、比較した結果、最も中心に近い領域が主要箇所であると判定するものであっても良い。

【選択図】 図2

出願人履歴情報

識別番号 [000000295]

1. 変更年月日 1990年 8月22日

[変更理由] 新規登録

住 所 東京都港区虎ノ門1丁目7番12号

氏 名 沖電気工業株式会社